

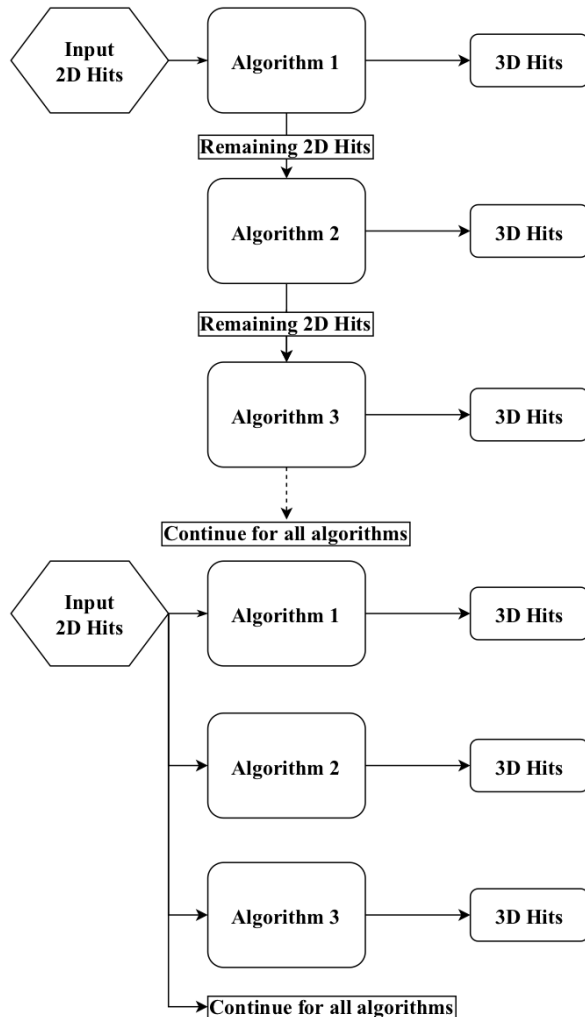
Improving the 3D spacepoint outputs from Pandora

Ryan Cross

DUNE FD Sim/Reco Workshop - 2020/09/14



Last Update



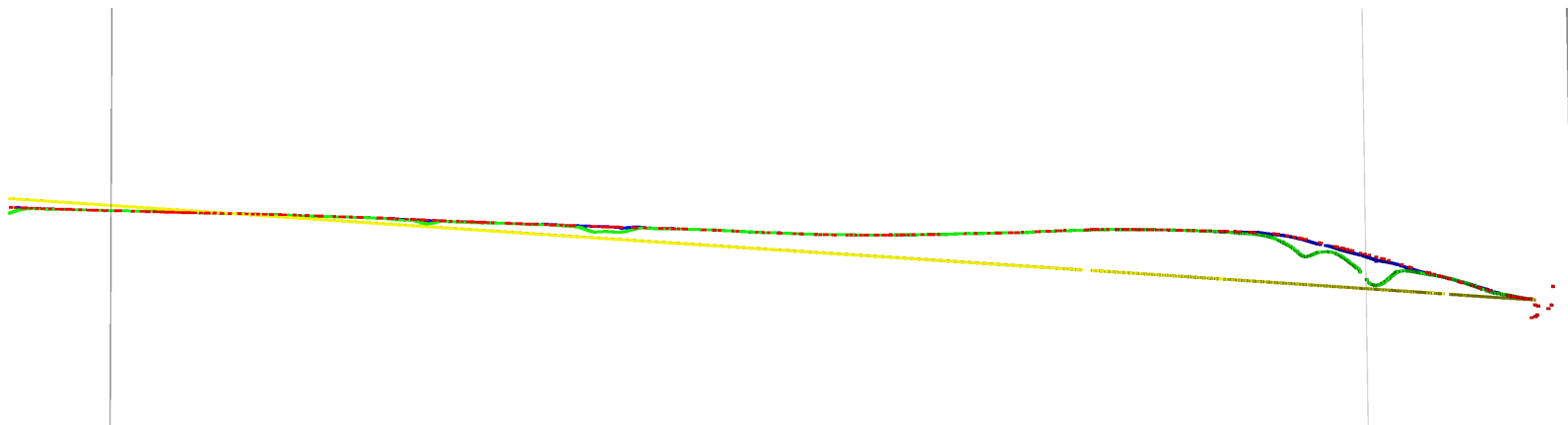
From the last update, I spoke about moving the 3D reconstruction tools (the part that combines 2D hits from each view into a single spacepoint) from a "waterfall" style model, to one that uses all tools equally, and then has to pick between the outputs in some fashion.

I have been updating this to produce many spacepoints, and then I can select the most consistent hits out of the large set of all hits. This avoids the waterfall approaches downside, where good tools are ignored if they are later in the process. All of this work is looking at track like particles in the DUNE FD, using MCC11 samples.

Changes Overview - The Inputs

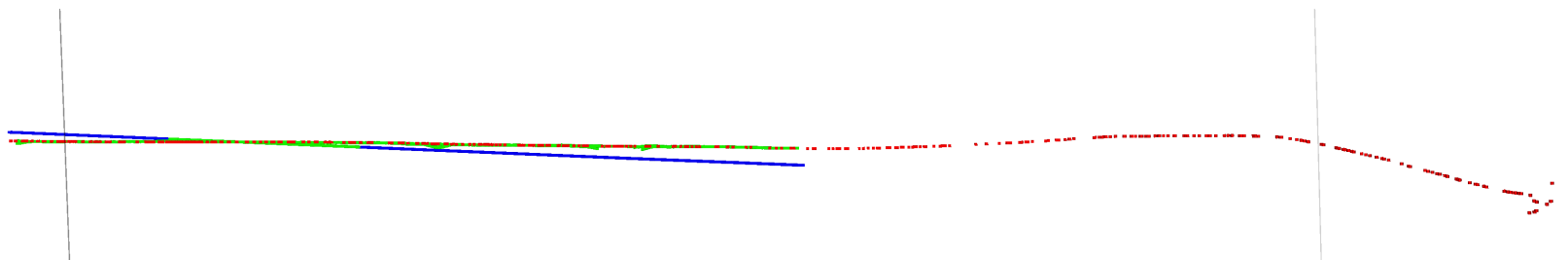
We start with the tools worth of outputs, which is 7 in the case of the standard configuration.

The dotted red in the image below represents the MC, every other colour is a different tools output.



Changes Overview - Initial Models

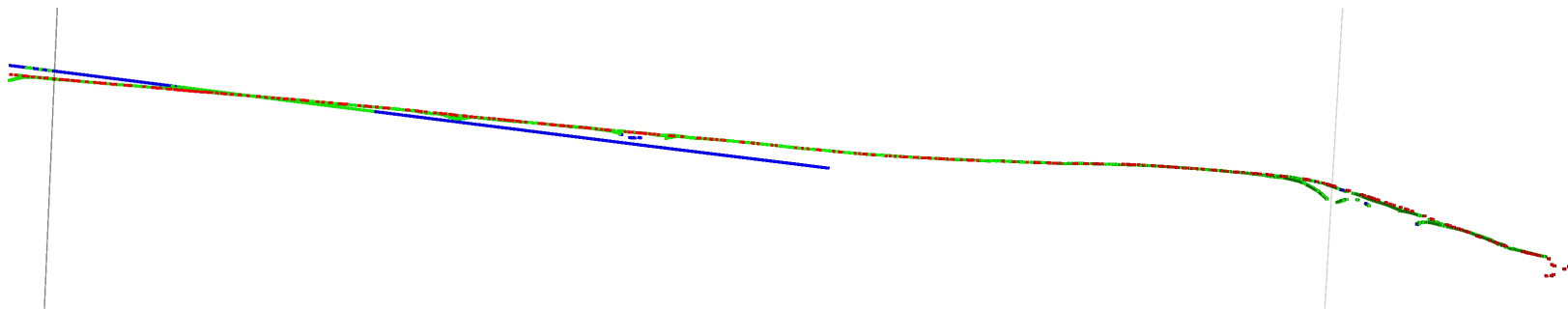
RANSAC (Random Sample Consensus) is used as a simple way to produce two simple straight line models that fit the data. Two are used to avoid cases where a single model would get stuck fitting errors/bad hits. The first fit is the model that best fits the data, and the second is the model that best fits the data, as well as being the most unique compared to the first model.



Changes Overview - Final Models

Once there is two models, they can be fixed up by evaluating all the remaining hits, and iteratively performing "fit extensions".

That is, using sliding linear fits, go over the chosen hits and look if there are any other hits nearby that look like they are consistent. This is done at either end, and lets the final result follow tracks that are not straight lines.

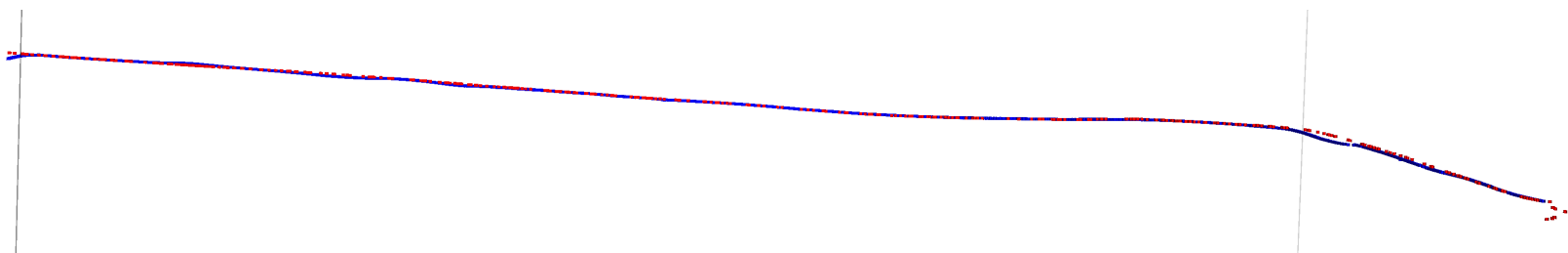


Changes Overview - Chosen Output

Next, the two complete models can be chosen between.

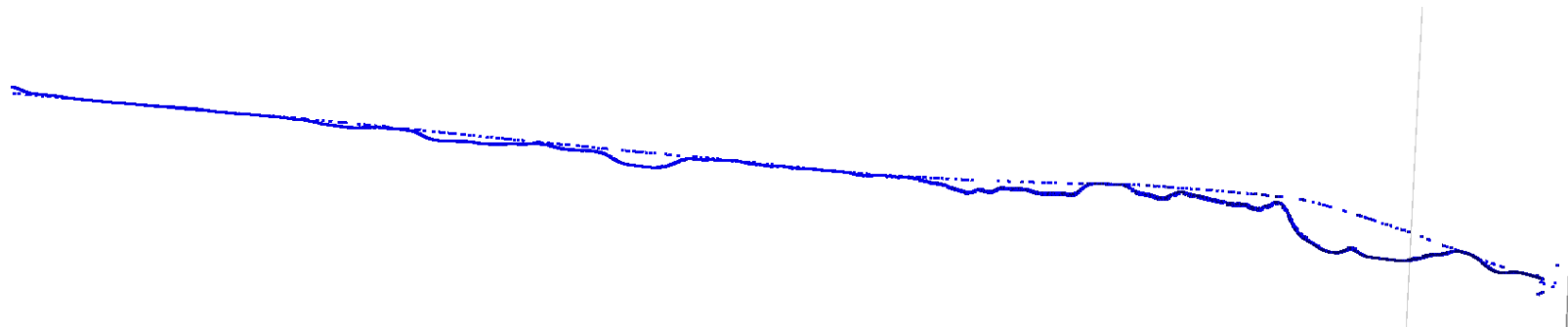
The decision is made using the number of spacepoints in the two final vectors, but also crucially using the count of hits that were considered in the previous step. This means the model that followed more spacepoints will be more preferable, which is desirable.

This chosen model goes through a final stage of clean up, which accounts for the changes between this output and the previous.

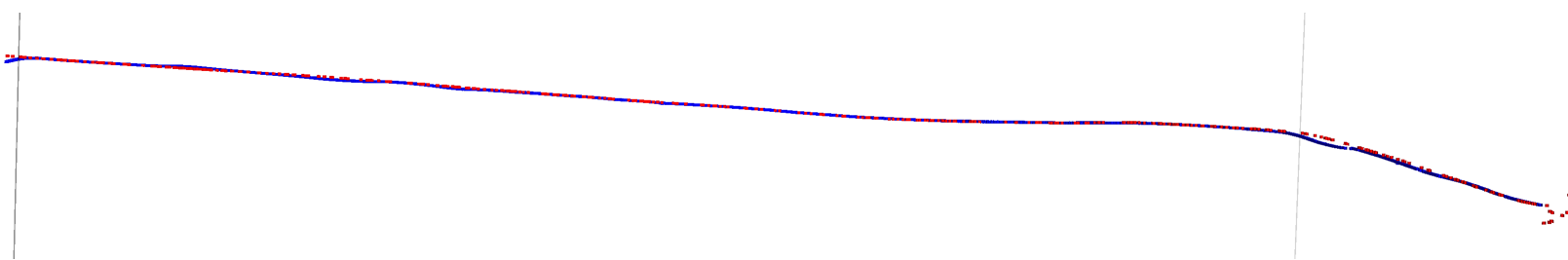


Changes Overview - Comparison

Before this work



After

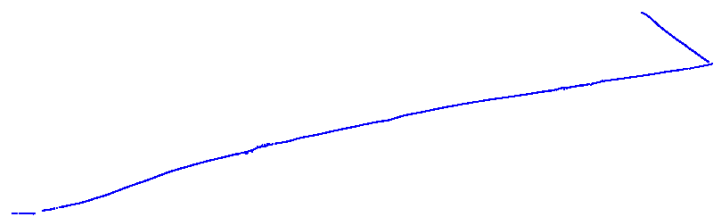
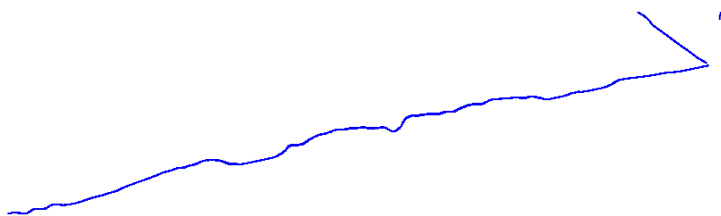


Performance Evaluation Outline

Measuring the performance of these changes has proved difficult, if only because when targeting a (luckily!) small number of particles, it can be easily lost especially when plotting per particle.

To help with that, I use a basic way of calculating "the event has changed this much", by plotting the event before and after my changes in a simple Python event display. I can then do low level mean squared error comparisons for the pixels themselves in the saved images. That gives a final score of how different they are, and is repeated from multiple angles.

This score is only used to select events to plot.



Performance Evaluation Outline

Measuring the performance of these changes has proved difficult, if only because when targeting a (luckily!) small number of particles, it can be easily lost especially when plotting per particle.

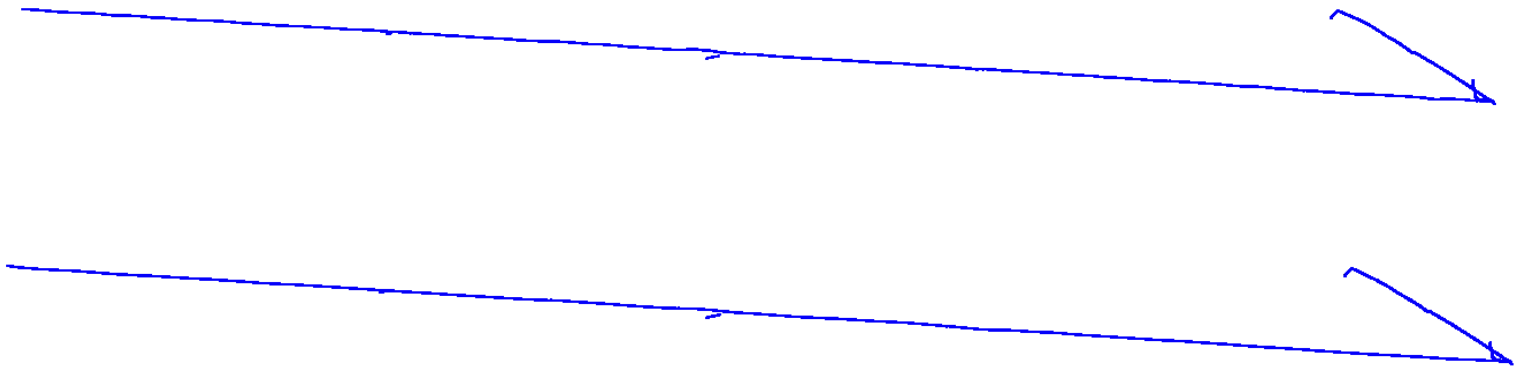
To help with that, I use a basic way of calculating "the event has changed this much", by plotting the event before and after my changes in a simple Python event display. I can then do low level mean squared error comparisons for the pixels themselves in the saved images. That gives a final score of how different they are, and is repeated from multiple angles.

This score is only used to select events to plot.



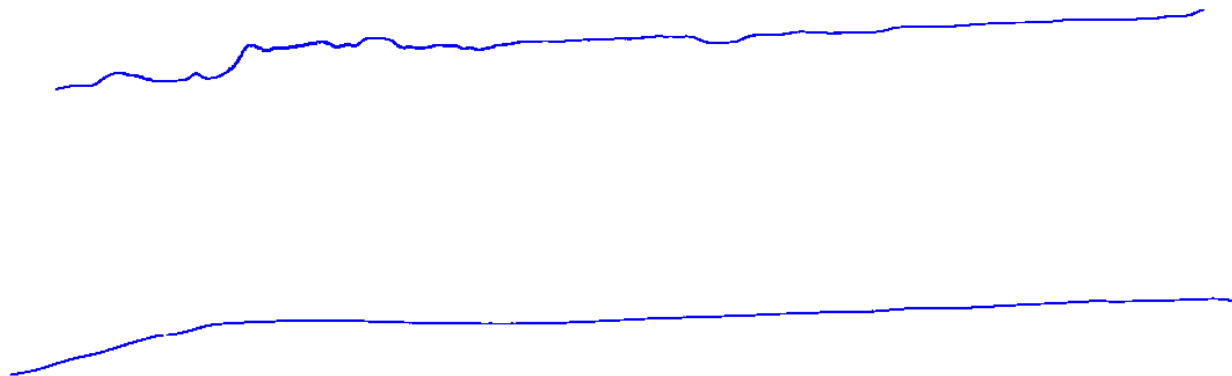
Example Low Score - 10.27

Since these scores are entirely arbitrary, here is an example of a small change.



Example High Score - 57.75

And a larger change.



Using the Score

This score is then used to select events to plot more in-depth results. A value of at least **40** is used for the following plots, which as shown in the previous slide, accounts for the large fixes that I am targeting.

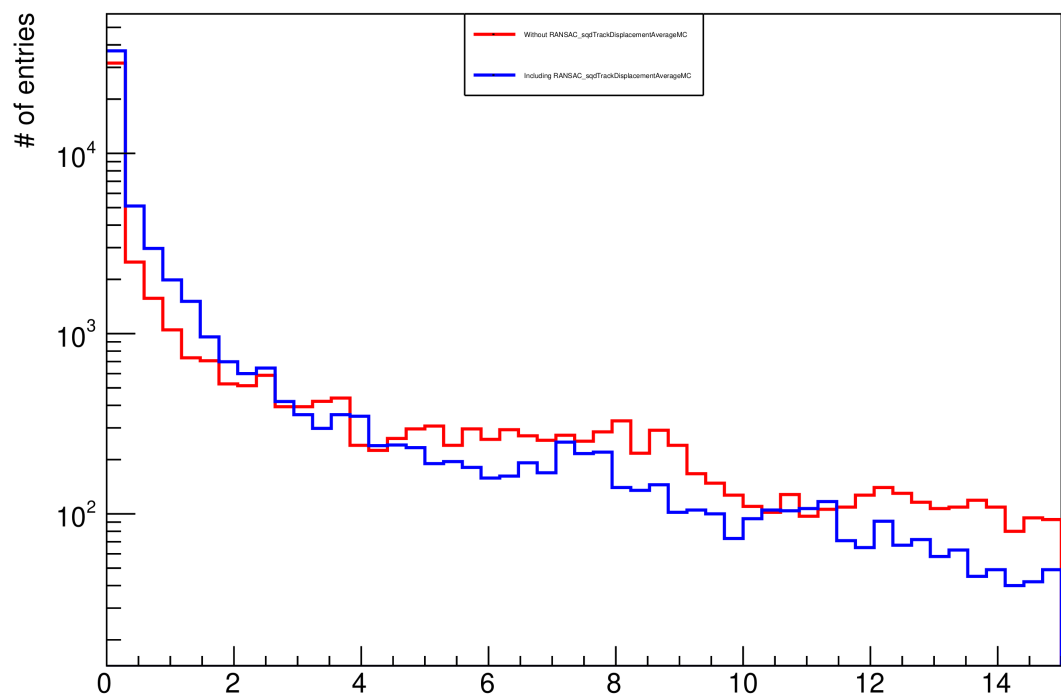
This score means that the various metrics I use (MC and Reco displacement, wiggle and ratio of 2D to 3D hits) can be targeted at the events that I am actually changing, and then the metrics can be used to see if the events that change are changing for the better.

This has the effect of meaning lower statistics as these events are rarer, but the changes for varying scores are included in a table as well.

As noted previously, the cut is done **per event**, and the metrics are **per particle**. This means there is always a mixture of particles, both ones that have changed and ones that have not, which means at lower scores the changes can get lost behind the other particles.

Squared MC Displacement

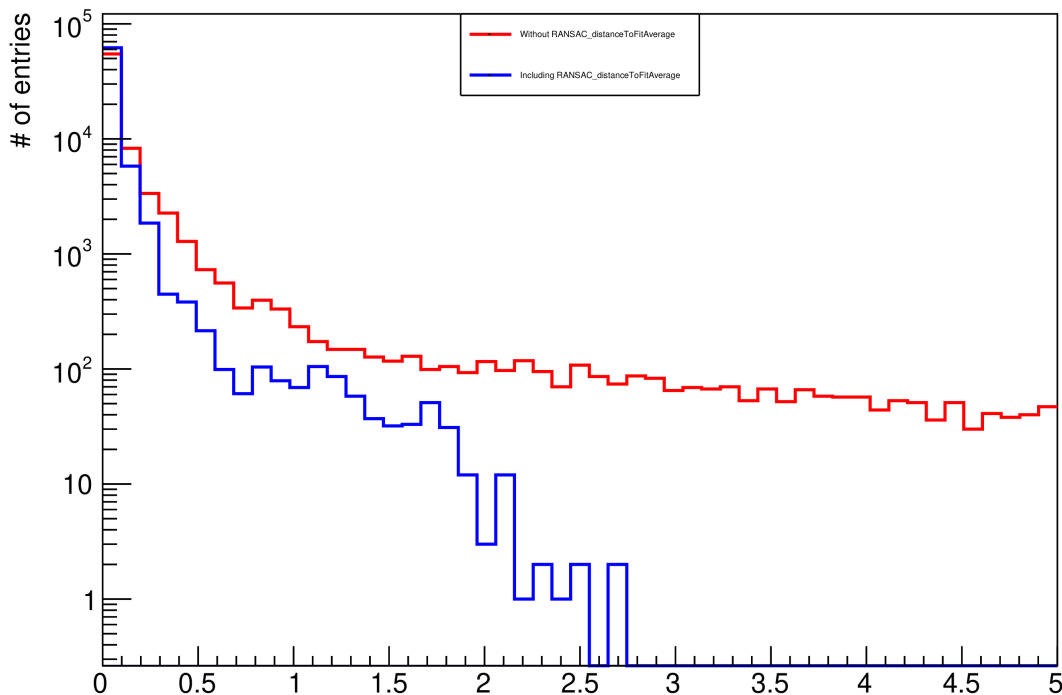
Direct MC comparisons are awkward in 3D, so instead, comparisons are made between two fits. The first is built on the MC and the second is built on the reco, and then their squared displacement can be compared at each hit to see how they differ.



Score	Sqd MC Disp Mean
40	0.64 -> 0.58 r
30	0.73 -> 0.61
15	0.70 -> 0.67
5	0.52 -> 0.55

Reco Displacement

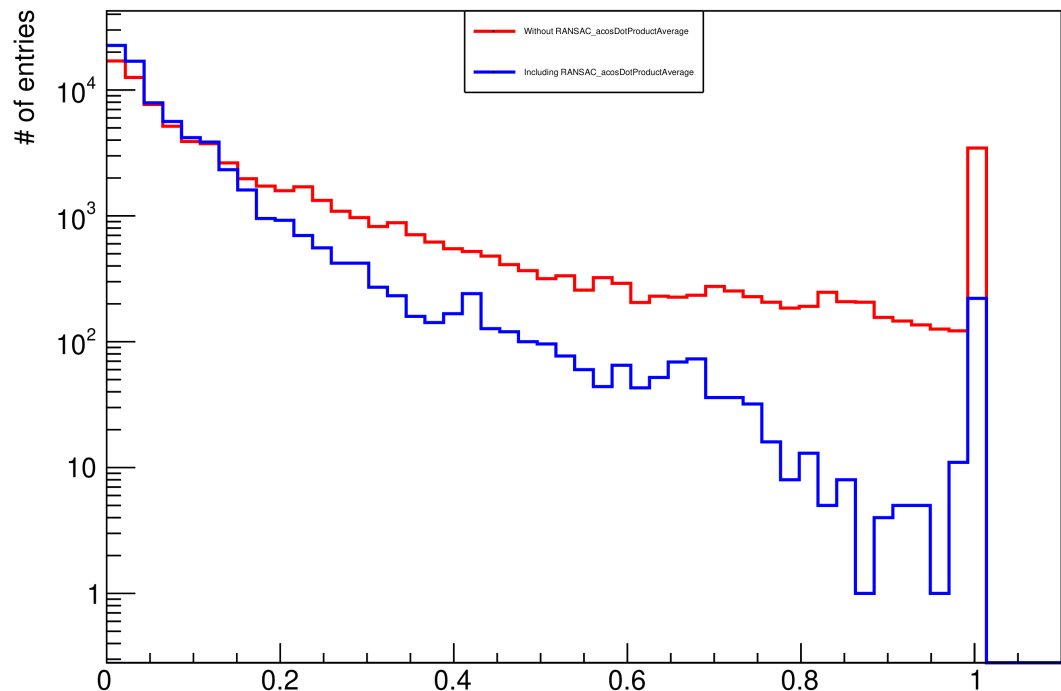
Similarly, we can calculate the displacement between each hit and a fit built over the full reconstructed track. This can somewhat be treated as an outlier check, how far away a hit is from the main fit. The displacement is calculated as $\frac{1}{3}(\Delta^2 + \Delta^2 + \Delta^2)$.



Score	Reco Disp Mean
40	0.44 -> 0.29
30	0.37 -> 0.28
15	0.34 -> 0.29
5	0.30 -> 0.28

"Wiggle"

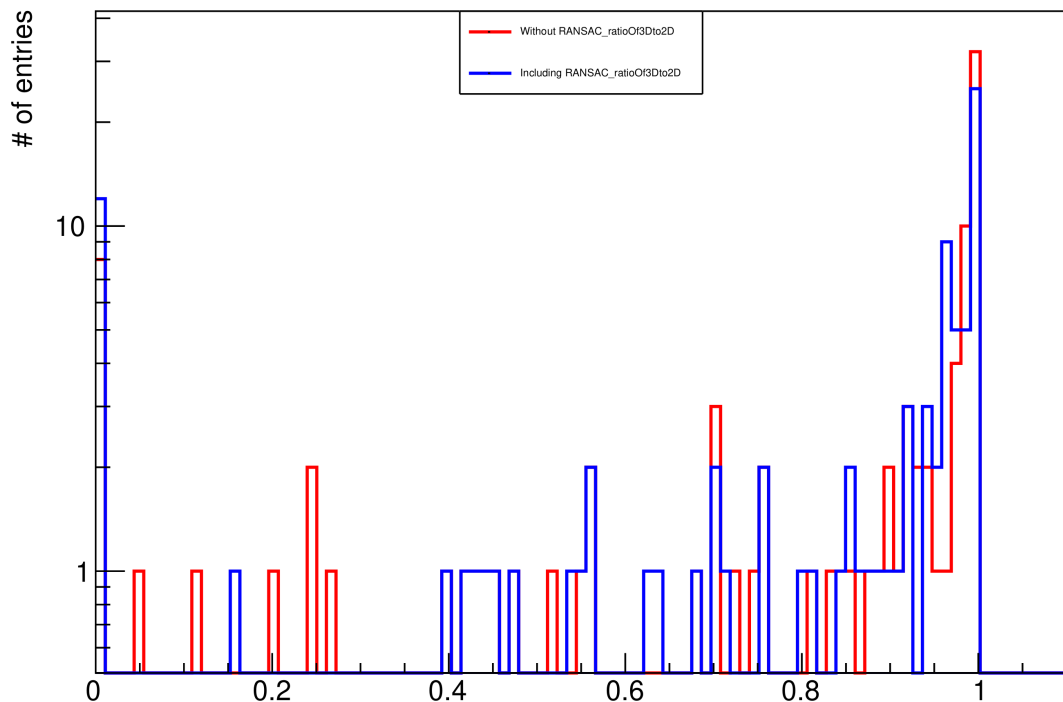
As previously seen, there can be a lot of wiggle in the produced tracks, so I calculate the dot product between the direction of the fit, and the direction that the current hit is moving relative to the fit.



Score	"Wiggle" Mean
40	0.12 -> 0.06
30	0.11 -> 0.05
15	0.10 -> 0.08
5	0.08 -> 0.07

Ratio of 2D to 3D

Finally, there is a few more plots that are less interesting plots that are nice to keep an eye on the performance, such as the average number of spacepoints, length of tracks and more. This shows the percentage of how many 2D hits have a corresponding spacepoint.



Score	Mean Ratio of 2D to 3D
40	0.77 -> 0.75
30	0.80 -> 0.79
15	0.83 -> 0.82
5	0.85 -> 0.84

Totals

In summary, the means are as follows:

Score	Sqd MC Disp	Reco Disp	AcosDotProd	Ratio of 2D to 3D
40	0.58 ± 0.17	0.029 ± 0.002	0.06 ± 0.01	0.75 ± 0.03
30	0.61 ± 0.12	0.028 ± 0.002	0.06 ± 0.01	0.79 ± 0.02
15	0.67 ± 0.08	0.029 ± 0.001	0.08 ± 0.01	0.82 ± 0.01
5	0.55 ± 0.04	0.028 ± 0.001	0.068 ± 0.004	0.838 ± 0.007

Issues

There is two main issues left:

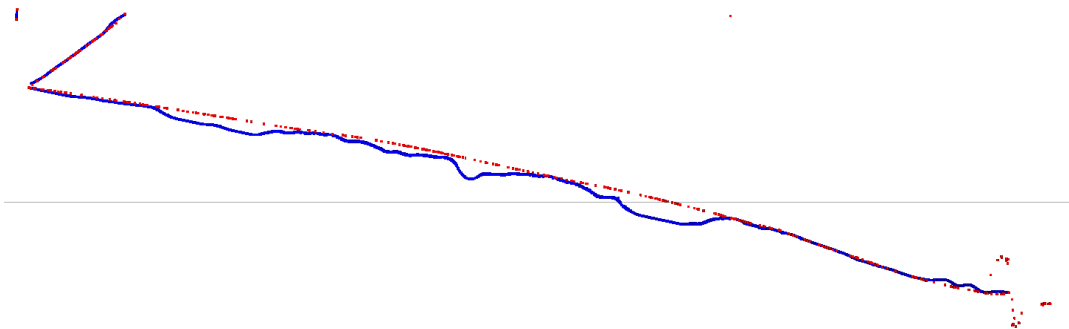
- There is no fundamental change in the spacepoint creation. That is, this is using the full power of all the existing tools. If the 2D reconstruction fails at some point, then this work will not correct or be able to help with that. This only ensures the best spacepoints are chosen.
- Slight drop in completeness due to growing. This is something that I believe can be worked out, as it is seemingly always due to insufficient fitting near the vertex. This makes sense as most work was initially focused on the fitting of the ends.

The first isn't really part of the scope of this work, but could be the target of some follow up work. However, from event scanning, I don't believe the issue lies with the 3D reconstruction itself any more, rather subtle issues in earlier stages causing issues. For example, if a delta ray is incorrectly marked as part of a track, this causes the 3D reconstruction to struggle much more due to increased ambiguities.

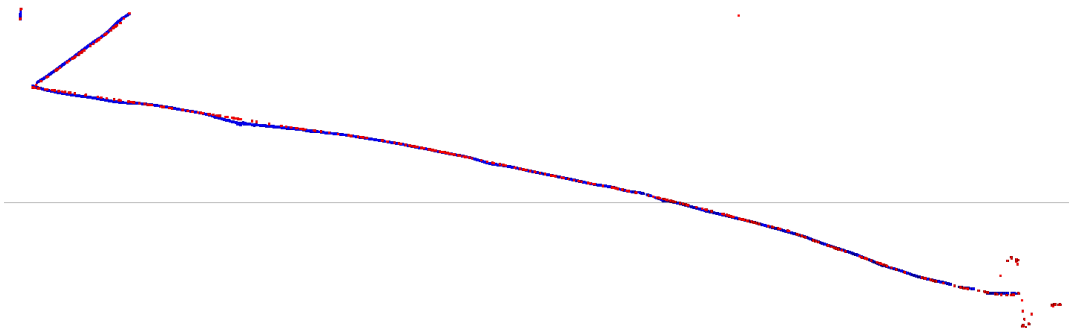
However, this approach does fix the worst events that are recoverable, but were failing due to the strict approach in the current Pandora release.

Example

Before this work

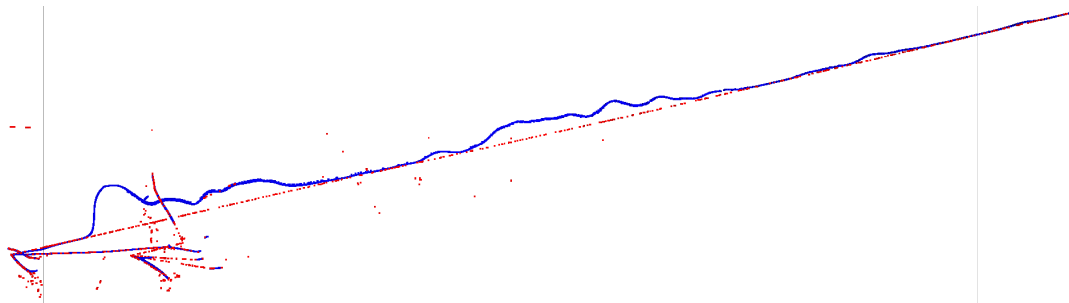


After

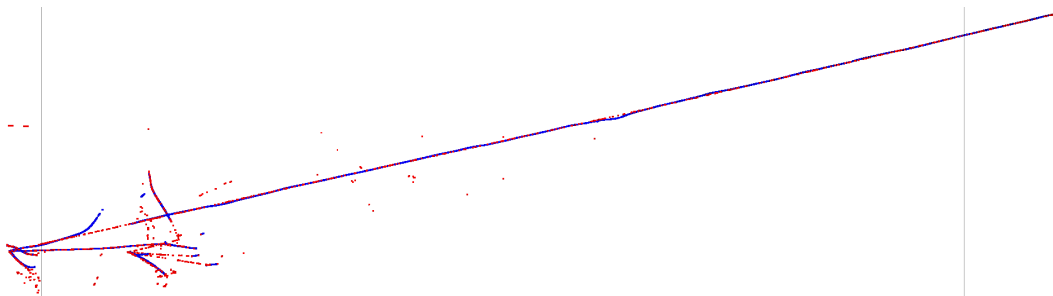


Example

Before this work

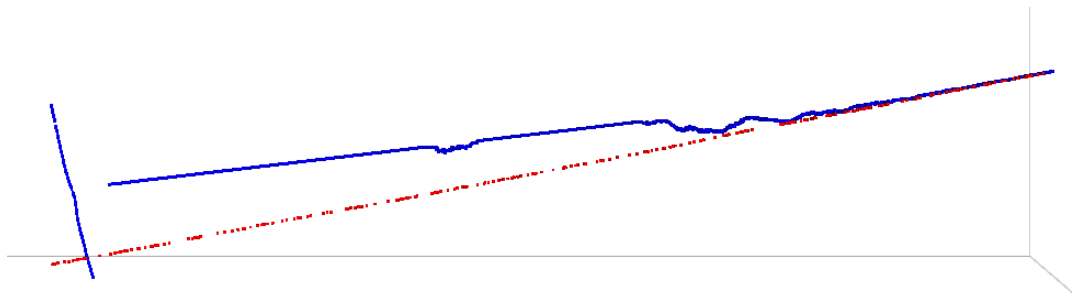


After

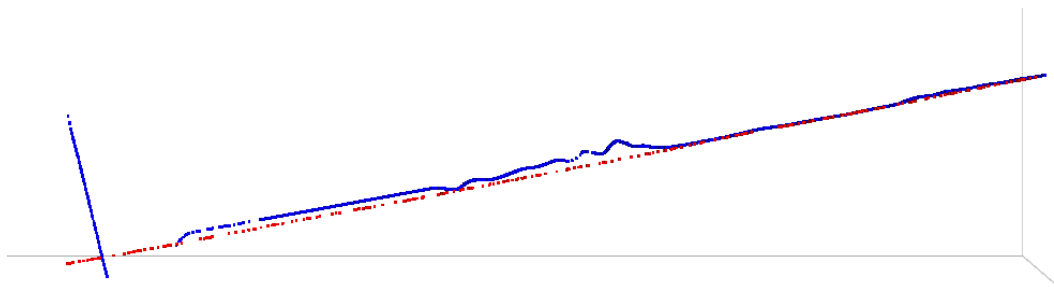


Example

Before this work

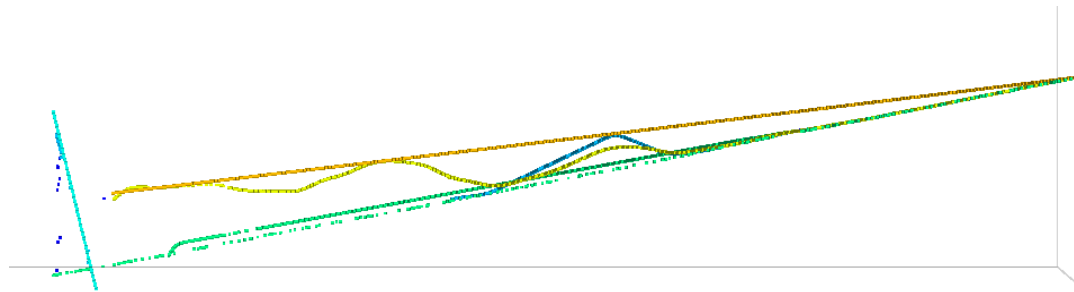


After

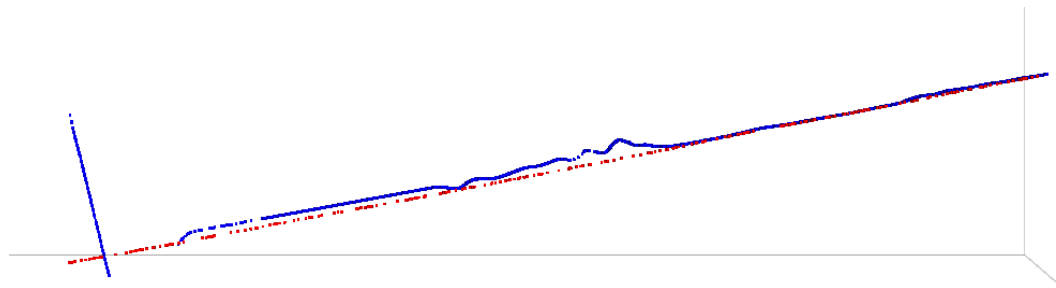


Example

Input Spacepoints



Result



Summary

To summarise, the spacepoint situation in Pandora is as follows:

- Previously, Pandora would use a strict ordering for the spacepoint creation, potentially causing issues in extreme cases.
- This has been refactored to always use every tool and create a superset of potential spacepoints.
- A process of selecting and improving two potential final outputs has been written, that tries to ensure that all the spacepoints produced are coherent and use the best produced spacepoints.
- With this update, most events are unchanged except those that had major issues previously, which should now be better. The few cases now that have issues are usually due to larger issues occurring earlier in the reconstruction chain.

This just needs a tidy up, and a final pass over some of the various thresholds, then it should be available as an opt-in feature.

Improving the 3D spacepoint outputs from Pandora

Ryan Cross

DUNE FD Sim/Reco Workshop - 2020/09/14

